Peripheral Blood DNA Methylation as Potential Biomarker of Malignant Pleural Mesothelioma in Asbestos-Exposed Subjects

Journal of Thoracic Oncology

Peripheral blood DNA methylation as potential biomarker of Malignant Pleural Mesothelioma in asbestos-exposed subjects.

--Manuscript Draft--

Manuscript Number:	JTO-D-18-00294R3
Full Title:	Peripheral blood DNA methylation as potential biomarker of Malignant Pleural Mesothelioma in asbestos-exposed subjects.
Article Type:	Original Article
Keywords:	Malignant pleural mesothelioma; DNA-methylation; case-control discrimination; Asbestos exposure; early biomarker
Corresponding Author:	Giuseppe Matullo, Prof Universita degli Studi di Torino Dipartimento di Scienze Mediche Torino, ITALY
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Universita degli Studi di Torino Dipartimento di Scienze Mediche
Corresponding Author's Secondary Institution:	
First Author:	Simonetta Guarrera
First Author Secondary Information:	
Order of Authors:	Simonetta Guarrera
	Clara VIberti
	Giovanni Cugliari
	Alessandra Allione
	Elisabetta Casalone
	Marta Betti
	Daniela Ferrante
	Anna Aspesi
	Caterina Casadio
	Federica Grosso
	Roberta Libener
	Ezio Piccolini
	Dario Mirabelli
	Irma Dianzani
	Corrado Magnani
	Giuseppe Matullo
Order of Authors Secondary Information:	
Manuscript Region of Origin:	ITALY
Abstract:	Introduction Malignant pleural mesothelioma (MPM) is an aggressive tumour strongly associated with asbestos exposure. Patients are usually diagnosed when current treatments have limited benefits, highlighting the need for non-invasive early diagnosis tests to monitor asbestos-exposed people. Methods

	We used a genome-wide methylation array to identify, in asbestos-exposed subjects, novel blood DNA methylation markers of MPM in 163 MPM cases and 137 cancer-free controls (82/68 Training Set; replication in 81/69, Test Set) sampled from the same areas. Results Evidence of differential methylation between MPM cases and controls was found (>800 CpG sites, Pfdr<0.05), mainly in immune system related genes. Considering the “top” differentially methylated signals, 7 single-CpGs and 5 genomic regions of coordinated methylation replicated with similar effect size in the Test Set (pfdr<0.05). The top hypomethylated single-CpG (cases vs controls effect size< 0.15, pfdr <0.05 in both Training and Test sets) was detected in FOXK1 (Forkhead-box K1) gene, an interactor of BAP1 which was found mutated in MPM tissue and as germline mutation in familial MPM. In the Test set, comparison of receiver operating characteristic (ROC) curves and the area under the curve (AUC) of two models, including/excluding methylation, showed a significant increase in case/control discrimination when considering DNA methylation together with asbestos exposure (AUC=0.81 vs AUC=0.89, DeLong’s test p=0.0013). Conclusions We identified signatures of differential methylation in DNA from whole blood between asbestos exposed MPM cases and controls. Our results provide the rationale to further investigate, in prospective studies, the potential use of blood DNA methylation profiles for the identification of early changes related to MPM carcinogenic process.

ORIGINAL ARTICLE

Peripheral blood DNA methylation as potential biomarker of Malignant Pleural Mesothelioma in asbestos-exposed subjects.

Simonetta Guarrera[a,b]*, Clara Viberti[a,b]*, Giovanni Cugliari[a,b], Alessandra Allione[a,b], Elisabetta Casalone[a,b], Marta Betti[c], Daniela Ferrante[d,e], Anna Aspesi[c], Caterina Casadio[f], Federica Grosso[g], Roberta Libener[h], Ezio Piccolini[i], Dario Mirabelli[j,k,l], Irma Dianzani[c,l], Corrado Magnani[d,e,l], Giuseppe Matullo[a,b,l,m]

[a] Italian Institute for Genomic Medicine, IIGM, Turin, Italy
[b] Department of Medical Sciences, University of Turin, Turin, Italy
[c] Department of Health Sciences, University of Piemonte Orientale, Novara, Italy
[d] Medical Statistics and Cancer Epidemiology Unit, Department of Translational Medicine, University of Piemonte Orientale, Novara, Italy
[e] Cancer Epidemiology Unit, CPO-Piemonte, Novara, Italy
[f] Thoracic Surgery Unit, AOU Maggiore Della Carità, Novara, Italy
[g] Division of Medical Oncology, SS. Antonio e Biagio General Hospital, Alessandria, Italy
[h] Pathology Unit, SS. Antonio e Biagio General Hospital, Alessandria, Italy
[i] Pneumology Unit, Santo Spirito Hospital, Casale Monferrato (AL), Italy
[j] Cancer Epidemiology Unit, Department of Medical Sciences, University of Turin, Turin, Italy
[k] Cancer Epidemiology Unit, CPO Piemonte, Turin, Italy
[l] Interdepartmental Center for Studies on Asbestos and Other Toxic Particulates "G. Scansetti", University of Turin, Turin, Italy
[m] Medical Genetics Unit, AOU Città della Salute e della Scienza, Turin, Italy

*S. Guarrera and C. Viberti contributed equally to this work

Short title: DNA methylation profiles as MPM biomarker

Corresponding author: Prof. Giuseppe Matullo, Department of Medical Sciences, University of Turin, Via Santena 19, Turin, 10126 Italy, and Italian Institute for Genomic Medicine, IIGM, Via Nizza 52, 10126 Turin, Italy; Phone: +39-011-6706501; Fax: +39-011-2365601; E-mail: giuseppe.matullo@unito.it

Funding

This work was supported by the AIRC - Associazione Italiana per la Ricerca sul Cancro [grant number 17464] to GM. This work is also supported by: the Compagnia di San Paolo, Turin, Italy, and the Italian Institute for Genomic Medicine (fka Human Genetics Foundation-Torino, HuGeF), Turin, Italy [grant 2016 to GM]; the Regione Piemonte, Ricerca Scientifica Applicata [grant to DM] for the Turin case-control study; the Asbestos Project grant of the Italian National Institute of Health [grant to CM]; grants from the Piedmont Region and Casale Local Health Authority [grant to CM] for the Casale Monferrato case-control study; the Ministero dell'Istruzione, dell'Università e della Ricerca – MIUR project "Dipartimenti di Eccellenza 2018 – 2022" [grant to GM]

Conflict of interests

CM and DM served as expert witnesses for the public prosecution office in asbestos–related court trials. All other authors declare no conflicts of interest.

Abstract

Introduction

Malignant pleural mesothelioma (MPM) is an aggressive tumour strongly associated with asbestos exposure. Patients are usually diagnosed when current treatments have limited benefits, highlighting the need for non-invasive early diagnosis tests to monitor asbestos-exposed people.

Methods

We used a genome-wide methylation array to identify, in asbestos-exposed subjects, novel blood DNA methylation markers of MPM in 163 MPM cases and 137 cancer-free controls (82/68 Training Set; replication in 81/69, Test Set) sampled from the same areas.

Results

Evidence of differential methylation between MPM cases and controls was found (>800 CpG sites, $_{Pfdr}$<0.05), mainly in immune system related genes.

Considering the "top" differentially methylated signals, 7 single-CpGs and 5 genomic regions of coordinated methylation replicated with similar effect size in the Test Set (p_{fdr}<0.05).

The top hypomethylated single-CpG (cases *vs* controls effect size<-0.15, p_{fdr} <0.05 in both Training and Test sets) was detected in *FOXK1* (Forkhead-box K1) gene, an interactor of BAP1 which was found mutated in MPM tissue and as germline mutation in familial MPM.

In the Test set, comparison of receiver operating characteristic (ROC) curves and the area under the curve (AUC) of two models, including/excluding methylation, showed a significant increase in case/control discrimination when considering DNA methylation together with asbestos exposure (AUC=0.81 *vs* AUC=0.89, DeLong's test p=0.0013).

Conclusions

We identified signatures of differential methylation in DNA from whole blood between asbestos exposed MPM cases and controls. Our results provide the rationale to further

investigate, in prospective studies, the potential use of blood DNA methylation profiles for the identification of early changes related to MPM carcinogenic process.

Keywords: Malignant Pleural Mesothelioma; DNA-methylation; case-control discrimination; asbestos exposure; early biomarker.

Abbreviations

MPM: Malignant Pleural Mesothelioma

CpGs: cytosine-guanine dinucleotides

DNAm: DNA methylation

DMe: differential methylation

DM-CpG: differentially methylated single CpG

DMR: differentially methylated region

WBCs: white blood cells

FDR: false discovery rate

PCs: principal components

ROC curves: receiver operating characteristic curves

AUC: area under the ROC curve

EEAA: extrinsic epigenetic age acceleration

IEAA: intrinsic epigenetic age acceleration

EWAS: epigenome-wide association study

Introduction

Malignant pleural mesothelioma (MPM) is an aggressive tumor strongly associated with asbestos exposure. It occurs long after the first exposure, with risk increasing depending on duration and intensity of exposure [1]. MPM incidence dramatically increased as a result of widespread use of asbestos in the past decades. Banning occupational asbestos use and preventing environmental asbestos exposure is rewarded as the only effective strategy to hinder MPM occurrence. However, for subjects who underwent environmental or occupational asbestos exposure in the past there is no preventive measure currently available, except for continuous monitoring aiming at an early MPM diagnosis, that might improve life expectancy. Patients are in fact usually diagnosed in an advanced phase, where radical surgery cannot be carried out, and the effectiveness of current cytotoxic therapies is limited [2], highlighting an urgent need for sensitive non-invasive tests for early MPM diagnosis.

In recent years, there is a growing interest in DNA methylation (DNAm) profile changes as possible diagnostic/prognostic biomarkers in cancer research [3].

DNAm is stable, yet it may be modified across lifetime by several factors including lifestyle, environmental exposures, ageing, and diseases [4] and is thus rewarded as an adaptive phenomenon potentially linking environmental factors and development of disease phenotypes. Being an early event in tumor development, aberrant DNAm has been suggested as a tool for early cancer detection and prognosis [5], and DNAm changes are regarded as possible actors also in MPM development and progression [6]. Along this line we sought whether white blood cells (WBCs) DNAm profiles might associate with MPM occurrence in subjects with asbestos exposure, considering that tumor environment in the pleura may trigger molecular changes in perfusing WBCs.

We thus investigated the occurrence of differences in the DNAm profiles in a large series of MPM cases and asbestos exposed controls, all of them with quantitative assessment of asbestos exposure.

Materials and Methods

Study population and asbestos exposure assessment

Study subjects belong to a wider ongoing collaborative study on MPM already described in previous reports. [7-9] Details on study population are provided in Supplementary Methods. Briefly, cases and controls were enrolled either in the municipalities of Casale Monferrato (Piedmont region, Italy), an area with high occupational and environmental asbestos exposure, or Turin (Piedmont, Italy).

Cases were enrolled in the main hospitals of the reference centers after histological confirmation of MPM diagnosis. Controls were randomly selected, from the local population (Casale Monferrato study) [7, 8], or among subjects not affected by neoplastic or respiratory conditions admitted to general medicine or urology units at the reference hospital (Turin study) [8].

For all the subjects cumulative asbestos exposure was quantitatively assessed as described in Ferrante *et al.*[10] Briefly, occupational history and lifestyle habits information was collected through interviewer-administered questionnaires filled out at enrolment during a face-to-face interview. An occupational epidemiologist assessed probability, frequency, intensity and duration of asbestos exposure (method detailed in in Ferrante *et al.* [10]) for both occupational and non-occupational exposures, taking into account their occupational and residential history and other possible circumstances of asbestos exposure, like sharing home with an asbestos worker.

In total, 300 samples, 163 MPM cases and 137 non-MPM controls were included in this study underwent DNAm analysis.

Our study complies with the Declaration of Helsinki principles, and conforms to ethical requirements. All volunteers signed an informed consent form at enrollment. The study protocol was approved by the Ethics Committee of the Italian Institute for Genomic Medicine (formerly Human Genetics Foundation, Turin, Italy).

DNA-methylation analyses

DNAm levels were measured in DNA from whole blood collected at subjects enrollment using the Infinium HumanMethylation450 BeadChip (Illumina) The percentage of methylation at each CpG site (methylation Beta-value), ranging from 0, no methylation, to 1, full methylation, was assessed for each subject. Laboratory methods for DNA extraction, BeadChip processing, methylation levels measurement, and data quality controls are detailed in Supplementary Methods.

Statistical methods

Study design is outlined in Supplementary Figure S1.

All statistical analyses were conducted using the open source software Rv3.4.0 [11].

Subjects were randomly allocated (R function *split*) to two analytical sets: a Training set of 82 MPM cases/68 controls (discovery panel), and a Test set of 81 cases/69 controls (replication panel).

Possible differences between the two sets were checked for sex, age, center, exposure, tumor histotype, smoking habits, WBCs estimated percentages. The same comparison was done for descriptive purposes within each one of the two sets (Table 1).

Epigenome-wide differential methylation between cases and controls was tested by regression models described below both at single CpG and at "regional" level. For all the regression analyses, continuous variables such as age (years), and asbestos exposure doses (fibers/mL-years), were Rank transformed (R function *rntransform*) to remove skewness. WBC subtype percentages from genome-wide methylation data for each subject were estimated [12] to evaluate differences in WBCs profiles between cases and controls. However, since numeric changes and functional dysregulation of immunocompetent cells triggered by asbestos and linked to MPM ethiopathogenesis were reported in literature (Table 2), regression analyses were not adjusted by WBCs composition (see Discussion section). Potential population stratification was taken into account by including in the regression analyses the first two principal components (SNP_PCs1-2) calculated on the basis of our previous genome-wide genotyping study [9] (Supplementary Methods); batch and technical effects correction was performed adjusting for "BeadChip control probes" PCs1-10 (CTR_Probes_PCs1-10; Suppl. Methods). Before inclusion as covariates into regression analyses, SNP_PCs1-2 and CTR_Probes_PCs1-10 variables were Rank transformed (R function *rntransform*). For multiple comparisons, Benjamini-Hochberg false discovery rate p (p_{fdr})≤0.05 was considered statistically significant.

Single CpG analysis

Cases-controls differential methylation (DMe) at Single-CpGs was tested in the Training set by generalized linear models analysis adjusted for asbestos exposure, gender, age, population stratification, and technical variability, as described above.

Smoking habit was purposely not added to our multivariate statistical model since smoking is not among the known risk factors for mesothelioma. However, since smoking habits may

modify DNA methylation profiles at some genomic locations, its effect on methylation levels was targetedly investigated on CpGs differentially methylated between cases and controls. To ensure a power of the study greater than 90% (two-tailed test, 0.05 alpha error), only CpGs with regression effect sizes differences between cases and controls $\geq|0.1|$ were considered in the discovery phase on the Training set, and subsequently investigated in the Test set under the same analytical conditions. CpGs with effect sizes differences between cases and controls $\geq|0.1|$ and $p_{fdr} \leq 0.05$ in both Training and Test sets were considered as replicating, and underwent gene set enrichment analysis (WEB-based GeneSeTAnaLysis Toolkit, WebGestalt; http://bioinfo.vanderbilt.edu/webgestalt).

Additionally, Test set subjects were clustered (unsupervised clustering analysis, R package *Rgplots*) according to their methylation levels of the differentially methylated single-CpGs (DM-CpGs). For this analysis, samples with missing methylation Beta-values for any of the CpGs were removed.

Finally, in the Test set, we compared the area under the receiver operating characteristic curve (AUC) of two models by the DeLong test : Model 1 included age, sex and asbestos exposure; Model 2 was as Model 1 plus methylation levels of significant and replicating DM-CpGs: this was done to test the potential improvement in case-control discrimination when considering DMe information together with asbestos exposure information.

Regional differential methylation analysis

We used the A-clustering algorithm to define genomic regions characterized by a group of two or more neighboring CpGs (CpGs clusters) with correlated methylation levels (Supplementary Methods) under the hypothesis that methylation changes in genomic regions may underlie potential functional changes linked to MPM. In the Training set, we tested the occurrence of differentially methylated regions (DMRs) between cases and controls by

generalized estimating equations with the same adjustments as the single-CpG analysis. DMRs with effect size differences of at least 0.1 between cases and controls were investigated for occurrence of DMe in the Test set under the same analytical conditions. DMRs replicating in the Test set (i.e. effect sizes $\geq|0.1|$ and $p_{fdr}\leq 0.05$ in both Training and Test sets) underwent gene set enrichment analysis.

CpGs methylation vs gene expression

Since our study envisaged the collection of DNA only, we were unable to assess RNA transcripts levels for our mesothelioma cases and related controls. To assess a possible relationship between differentially methylated CpGs and gene expression variations of the corresponding genes, we took advantage of other in-house studies for which whole transcriptome gene expression and DNAm data were available. Data from two sample sets were used: i) 72 healthy subjects included in a genotype-phenotype correlation study (from here on, Geno-Pheno) described elsewhere [13] for which methylation and gene expression were assessed on DNA and RNA from the same aliquot of peripheral blood mononuclear cells (PBMCs), and ii) 21 normal pleura surgical samples from donors that were subjected to thoracoscopy for conditions different from MPM, belonging to a previous study [8], for which whole genome methylation and whole transcriptome expression were recently assessed as a pilot study (data not shown). In the two studies, methylation levels were determined by the HumanMethylation450 BeadChip and quality controlled as described in Supplementary Methods. Gene expression levels were determined by the Illumina HumanHT-12 v4 Expression BeadChip according to manufacturer's instructions (Supplementary Methods). Genes were considered expressed if their probe(s) intensity signal(s) was(were) at least twice the BeadChip background. The relationship between methylation and transcript levels of genes was investigated with a Pearson's correlation test.

DNA Methylation Age estimation and Age Acceleration Indices

DNA Methylation age (DNAm-age) was estimated for each subject from DNAm levels on the overall sample according to the method proposed by Horvath *et al.* [14].

Estimated DNAm-age was compared with chronological age by the Spearman correlation test. Age acceleration (AA) was calculated as the residuals of the DNAm-age estimate regressed on chronological age in both MPM cases and cancer-free controls: positive AA suggests a biological age "older" than chronological age, whilst a negative one suggests the individual is biologically "younger".

Two additional age acceleration indices, intrinsic epigenetic age acceleration (IEAA) and extrinsic epigenetic age acceleration (EEAA), were calculated as well. As reported by Chen *et al*. [14], EEAA is influenced by blood cell count contribution, whereas IEAA value is only weakly correlated with estimated measures of blood cell counts. The two indices, estimated from DNAm data, are thus indicators of 2 different cellular ageing processes, one (EEAA) dependent on and the other (IEAA) independent from WBC count. Differences between cases and controls were tested by W test with $p \leq 0.05$ considered statistically significant, and by a generalized linear models analysis adjusted for sex and asbestos exposure levels.

Results

Sample descriptive statistics are reported in Table 1.

The distribution of subjects' characteristics of cases and controls included in the Training set and cases and controls included in Test set is rather homogeneous and comparable: Training and Test set subjects did not differ by age, tumor histology, smoking habits, exposure dose and source, estimated WBCs percentages (W test), sex, center of recruitment, case-control

distribution (Chi-square Test), except for a difference in B cells distribution in controls between Training set and Test set (W test, p=0.04).

Conversely, when comparing cases *vs.* controls within each of the 2 sets, cases were exposed to significantly higher cumulative doses of asbestos than controls, and had statistically significant differences in WBCs distribution with respect to controls (Table 1).

After quality controls on raw DNAm data, 389,147 CpG were used in the following analyses.

Single-CpGs differential methylation analysis

In the Training set, 887 CpGs had effect size ≥|0.10| at case-control DMe analysis, 884 of them with p_{fdr}<0.05. Thirteen CpGs were hypermethed in cases, while 871 were hypomethylated (Supplementary Table S1). Out of the 884 significantly differentially methylated CpGs in the Training set, 868 showed significant DMe ($p_{fdr\ targeted}$<0.05) also in the Test set with concordant and similar effect sizes (Supplementary Table S1), and were annotated to 599 genes that were analyzed for pathway enrichment. We found statistically significant enrichment for Neutrophil Degranulation, Innate Immune System and Immune System (Supplementary Table S2-A).

Appraised the huge number of hypomethylated signals (N=871) identified in the Training set, we posed a more stringent threshold at effect size ≤-0.15, obtaining a smaller group of 20 hypomethylated CpGs. The resulting DM-CpGs list from the discovery analysis in the Training set, thus, included 20 hypomethylated CpGs with effect size ≤-0.15 and 13 hypermethylated with effect size ≥0.10. Checking for strictly replicating signals in the Test set, we ended up with 7 "top DM-CpGs", 3 hypomethylated CpGs with effect sizes ≤-0.15 in both Training and Test sets (genes *FOXK1*, *MYB*, and *TAF4*) and 4 CpGs hypermethylated with effect sizes ≥0.10 in both Training and Test sets (genes *CXCR6/FYCO1*, *TAP1*, *MORC2*, *LIME1*) (Table 3).

We then performed two unsupervised clustering analyses on Test set samples to inspect the distribution of cases and controls according to their methylation levels at the DM-CpGs identified in the Training set: the first one (Supplementary Fig. S2-A) included the top DM-CpGs, irrespective of strict replication criteria, i.e. all the 33 top DM-CpGs, identified in the Training set; the second one (Supplementary Fig. S2-B) included only the 7 DM-CpGs whose effect sizes strictly replicated in the Test set, as described above. Cutting the sample dendrograms at the first node, in both analyses the groups with the most hypomethylated/hypermethylated DM-CpGs included the highest proportion of cases as compared to the other group (Supplementary Fig. S2-A,B).

The AUC comparison of 2 models (see methods) showed a statistically significant improvement in discrimination between cases and controls when including methylation levels into the analysis (Figure 1).

A univariate regression analysis in the overall MPM cases (N=163) was done to check for a possible relationship between methylation levels at each of the 7 DM-CpGs vs. tumor histotype, coded as epithelioid, sarcomatoid/fibrous, biphasic/mixed, non-defined, and unknown. Taking "epithelioid" histotype as the reference one, no clear evidence of methylation differences linked to different histotypes was found, except for the methylation levels of cg08450017 in *CXCR6/FYCO1*, that were significantly slightly reduced in biphasic/mixed MPMs with respect to epithelioid type (effect size±s.e.=-0.05±0.03, p=0.048).

To assess if smoking habits could modify DNA methylation profiles at the 7 DM-CpGs, we categorized all subjects as current- former- and never-smokers, and then performed a multivariate regression analysis in the overall sample (163 MPM cases,137 controls) with the same model used for the "discovery phase" adding also smoking status as covariate. No evidence of methylation differences linked to different smoking categories was found for any of the "top" CpGs, except for the methylation levels of cg00446123 in *LIME1*, that were

significantly slightly reduced in ex-smokers with respect to current smokers taken as the reference group (effect size±s.e.=-0.034±0.017, p=0.049).

For a subset of cases, for which more detailed clinical information was available, we checked for a possible relationship between methylation levels of the "top" 7 CpGs and time of blood drawing, and cTNM. Results are reported in Supplementary File (Supplementary Results).

Lastly, the possible relationship between methylation levels and genomic variability was checked for *FOXK1*, that is located in region 7p22.2 which was found associated with MPM in our previous GWAS report [9] and in the companion paper of our collaborators [15]. Results are reported in Supplementary File (Supplementary Results).

Regional differential methylation analysis

The A-clustering algorithm identified 24,573 CpGs clusters with correlated methylation levels in the Training set and 23,676 in the Test set, that were tested for differential methylation between cases and controls. As a first step, we checked for clusters of differentially methylated CpGs around the 7 "top DM-CpGs" identified at the single-CpGs analysis, and found that 5 out of the 7 top DM-CpGs identified at the single CpGs analysis were indeed included in CpGs clusters differentially methylated between cases and controls (Table 4).

As a second step, across the 24,573 CpGs clusters identified in the Training set, we looked for DMRs between cases and controls with effect size $\geq|0.10|$, and found 35 DMRs hypomethylated ($p_{fdr}<0.05$) in Training set cases. Among the 35 regions, the broadest cluster included 9 CpGs in *FOXK1* gene around cg04572930, that was the most hypomethylated CpG identified at single-CpG analysis. No region was hypermethylated above the established threshold (effect size $\geq|0.10|$) in cases with respect to controls in the Training set.

All the 35 regions identified in the Training set were represented as clusters also in the Test set, and all of them showed statistically significant hypomethylation in cases (range[-0.12;-0.07], median=-0.10, p_{fdr}<0.05) (Supplementary Table S3). Restricting the analysis to clusters with effect size ≤-0.10 and p_{fdr}<0.05 in both Training and Test sets (i.e. strictly replicating DMRs), 21 regions corresponding to 21 genes were identified. The 5 regions with more than 2 CpGs in cluster were noted as "top DMRs", and corresponded to the following genes: *FOXK1* (9 CpGs), *CSTA* (4 CpGs), *ZNF516* (3 CpGs), *TOLLIP* (3 CpGs), *TNFAIP6* (3 CpGs).

At a gene-set analysis on the 21 DMRs, no statistically significant enriched category rwas found. However, considering enriched categories that included more than 1 gene, we found the most represented pathways were related to immune systems processes (3 categories, 5 genes in multiple categories), cancer related pathways (5 categories, 7 genes in multiple categories), and developmental biology (1 category, 5 genes) (Supplementary Table S2-B).

CpGs methylation vs Gene expression

The methylation levels of the 7 differentially methylated single-CpGs and of the 21 CpGs from the regional analysis were correlated with the corresponding gene-expression levels in 2 separate datasets: one comprising methylation and gene expression data from healthy subjects (Geno-Pheno) PBMCs', and the other one comprising methylation and gene expression data from healthy pleura.

In the Geno-Pheno dataset we found no significant correlation between methylation levels and transcript levels in PBMCs for any of the considered CpGs.

In the healthy pleura dataset, cg08450017 in *FYCO1*, and 8 out of 9 CpGs in the *FOXK1* cluster showed a statistical significant correlation between methylation levels and the corresponding gene transcript levels (Table 5).

DNA Methylation Age and Age Acceleration

DNAm-age was estimated for each subject from methylation levels in blood cells, and showed a strong correlation with chronological age (Spearman's rank correlation rho=0.80, p<2.2E-16).

AA index did not significantly differ between cases and controls: Wilcoxon rank sum test, residuals median[interquartile range], AA_{cases}=176.0[97.5;256.0], $AA_{controls}$=168[65.0;261.0], Wilcoxon p=0.50; GLM analysis, effect size(se)= -2.24(12.81), p=0.86.

Considering the additional AA indices estimated on the overall sample, IEAA, which accounts for WBCs composition, did not significantly differ between cases and controls: median[interquartile range], $IEAA_{cases}$=0.22[-3.17;3.37], $IEAA_{controls}$=0.06[-3.82;2.80], Wilcoxon p=0.49, GLM effect size(se)= 0.74(0.68), p=0.28.

On the other hand, EEAA, which does not account for WBCs composition, was significantly different between cases and controls: $EEAA_{cases}$=0.51[-3.35;5.75], $EEAA_{controls}$=-0.42[-5.04;2.57], Wilcoxon p=0.007, LM effect size(se)= 2.27(0.86), p=0.009.

Discussion

Altered DNA methylation is frequently observed in cancer, and DNAm profiles of specific genes were already proposed as potential tools for cancer detection, risk prediction, and prognosis [5]. In our study we investigated WBCs DNAm profiles in a large series of MPM cases and controls, all of them with quantitative assessment of asbestos exposure, with the aim of identifying molecular hallmarks of MPM in non-invasively collected blood samples from asbestos-exposed subjects. The role of the immune system in cancer is well known , including mesothelioma [16], and there are evidences that asbestos related antigenic overstimulation and ROS oxidation trigger functional changes in WBCs [17]. Moreover, several

previous reports, summarized in Table 2, suggest asbestos-induced immune system deregulation, autoimmune response and functional changes in WBCs, putatively linked to MPM occurrence.

Indeed, the reduction of estimated CD4+ and CD8+ T lymphocytes in MPM cases, that we observed in our study, suggests a weaker adaptive immune system, and is compatible with the possible occurrence of functional changes in cellular subpopulations in MPM [18].

Considering these evidences, we hypothesized that both numerical and functional changes in WBCs might be reflected in methylation changes in MPM cases compared with controls, thus we did not include WBC estimated counts as covariate in our analyses. In general, WBCs model adjustment is meant to control for inter-individual variability, which might be essential when healthy subjects are compared, but might introduce biases when analyzing subjects with diseases that trigger immune system response resulting in different cell counts and activation profiles.

In our sample MPM cases showed "older" WBCs than controls according to epigenetic ageing profiles: the EEAA index [14], which does not account for differences in WBCs compositions, is significantly different between cases and controls, indicating that cases' WBCs are epigenetically older than those of controls. Epigenetic age acceleration was associated with mortality and age-related diseases [19-21], and may play a role in MPM too, making it worthy of further investigation.

In this study we identified signals of cases-controls differential DNAm both at single-CpG and regional analyses. The "top" hypomethylated DM-CpG is located in *FOXK1* (Forkhead Box K1), a transcription factor involved in development and metabolism, and Wnt signaling[22]. We found *FOXK1* differentially methylated also at regional analysis, with a 9 CpGs cluster hypomethylated in cases. The involvement of FOXK1 in tumor onset and progression was previously reported for colorectal and gastric cancers [23, 24]. It was also recently reported that

mTORC mediated de-phosphorylation of FOXK1 transactivates *CCL2* gene promoting tumor associated macrophages infiltration in mice, contributing to tumor progression. [25]

Moreover, FOXK1 directly interacts with BAP1 (BRCA1-associated protein 1) [26], whose involvement in MPM has been already described [27-29].

Notably, *FOXK1* is located on chromosome 7p22.2, one of the 5 genomic regions that we previously identified as associated to MPM in a GWAS on Italian samples and that were further replicated in an independent Australian MPM sample [9]. However, no evident relationship between methylation levels and genomic variation at *FOXK1* locus was found, at least for the 6 SNPs for which the genotypes were available in our sample (Supplementary File, Supplementary Results).

The most hypermethylated single-CpG signal is located in *FYCO1* (FYVE And Coiled-Coil Domain Containing 1) gene body, and in the promoter region of *CXCR6* (C-X-C Motif Chemokine Receptor 6) gene, which is antisense transcribed to *FYCO1* intron 14.

Interestingly enough, *FYCO1* encodes for a putative FOXK1 interactor within the human autophagy network [30]. In the present study, CpGs in *FYCO1* and *FOXK1* are the only signals whose methylation levels showed significant positive correlation with gene expression levels, as measured in a set of healthy pleura samples.

CXCR6, a T cell chemokine receptor, and its ligand, the chemokine CXCL16, were recently described as markers and promoters of inflammation-associated cancers, such as prostate cancer, and were suggested as mediators of inflammation-related tumorigenesis through direct effects on cancer cell growth and by inducing the migration and proliferation of tumor-associated leukocytes [31].The same authors reported that the expression of CXCR6/CXCL16 in both cancer cells and adjacent T cells correlated with prostate cancer progression.

Out of the 868 DM-CpGs (599 genes) and the 21 DMRs (21 genes), we found several signals in genes belonging to “Neutrophil Degranulation”, “Innate Immune System”, and “Immune

System" Reactome pathways. This is not surprising, since we are studying DNAm in WBCs with the purpose to identify relevant changes potentially associated with processes leading to MPM onset.

Among the "top signals", three out of the 5 hypermethylated DM-CpG i.e. *TAP1* (Transporter 1, ATP Binding Cassette Subfamily B Member), *LIME1* (Lck Interacting Transmembrane Adaptor 1) and *CXCR6* (C-X-C Motif Chemokine Receptor 6), as well as 3 of the top DMRs, i.e. *CSTA* (Cystatin A), *TOLLIP* (Toll Interacting Protein) and *TNFAIP6* (TNF Alpha Induced Protein 6) have a role in the immune system and/or inflammation-related processes. [32-35].

Taken together, our findings of DMe in WBCs suggest that DNAm changes in the immune system components may play a role in or may be a consequence of the oncogenic process triggered by asbestos exposure. However, a major limitation to the functional interpretation of our results is that all MPM subjects have already developed the disease at recruitment: thus, our findings likely reflect the disease status rather than being causal factors in the dynamic processes leading to MPM onset. Cases were enrolled at MPM diagnosis, when the tumor was already in place, so we cannot exclude a "reverse causality" bias, and we are thus unable to establish whether methylation changes are associated with the carcinogenic process, or are consequence of the disease.

Another limitation is the unavailability of MPM tissue from the same subjects, that also poses major constrains to the functional interpretation of our findings and prevents the possibility to compare DNAm profiles in blood cells and in tissue samples or pleural effusions of MPM cases. Even though we could not prove, in the context of this study, whether or not DNA methylation profiles in blood reflect the status of the tumor, in our view they may be assumed as "biomarkers" of actual disease (given the cross-sectional nature of the study). As a matter of fact, notwithstanding the above limitations, the discrimination between MPM cases and

non-MPM asbestos exposed controls improved when DNAm levels were taken into account together with asbestos exposure levels.

The significant difference in degree of asbestos exposure between case and control is a critical issue in verifying these signatures as markers for the early detection of MPM in asbestos-exposed subjects. For this reason, we corrected our analyses for individual levels of asbestos exposure, taking advantage of well recorded asbestos exposure information, and thus we identified DNA methylation biomarkers which are not directly associated to asbestos exposure. Since it is not always possible to accurately quantify asbestos doses for exposed subjects; it is important to develop tools to confidently identify those asbestos exposed subjects who are prone to MPM occurrence also when it is not possible to assess the cumulative dose they have been exposed to. The here reported evidences of detectable methylation differences in asbestos exposed subjects with and without MPM provide the rationale to search for possible DNA methylation early changes associated to MPM development. Monitoring methylation changes on the trajectory to the disease, from asbestos exposure to overt MPM diagnosis, is a necessary task for the identification of early epigenetic changes biomarkers related to MPM to be employed for the monitoring of asbestos-exposed subjects. MPM is characterized by the accumulation and persistence of asbestos fibers in the lungs, leading to a long latency period before clear clinical signs of the tumor are detectable. As of today, early MPM detection is still poor. Serum biomarkers such as Mesothelin, megakaryocyte potentiating factors, Osteopontin, CYFRA21-1, and Fibulin-3 were tested for early MPM diagnosis with limited results [36] and, so far, neither radiologic nor biochemical screening studies have proven useful. In our view, progressive changes in WBCs' DNA methylation levels might thus provide a tool to monitor progressive alterations in the immune system of asbestos exposed subjects, potentially related to MPM development (**Figure 2**). Further studies in prospectively collected cohorts of asbestos-exposed subjects with blood

sampling at multiple points across time will be needed to clarify if, and to what extent, methylation changes progress across time towards the tumor onset. This is particularly important since, despite the banishment of asbestos itself and asbestos-containing materials in several countries, individuals who have already been exposed will remain at risk of MPM, whose incidence is just expected to peak in the next years in European countries.

Overall, our study identified signatures of differential methylation in DNA from whole blood between asbestos exposed MPM cases and non-MPM controls, both at single-CpGs and at genomic region levels. We suggest that differential methylation patterns of selected CpGs in DNA from WBCs may aid in discriminating MPM cases from asbestos-exposed non MPM subjects and can be used to further improve MPM risk estimation is subjects occupationally and/or environmentally exposed to asbestos.

Figure1 legend

Figure 1: Receiver Operating Characteristic (ROC) curves on Test set

Model 1 (dotted line): age, sex, asbestos exposure levels

Model 2 (solid line): age, sex, asbestos exposure levels, CpGs methylation levels

Figure 2: hypotheses generation and rationale for future studies following the identification of differential methylation between asbestos exposed subjects with and without MPM.

Additional Material

Supplementary File.docx: Supplementary File

Supplementary Figure S1.pdf: Study design outline

Supplementary Figure S2.pdf: Clustering Analysis

Supplementary Table S1.pdf: Single-CpGs differential methylation analysis tables

Supplementary Table S1.docx: WebGestalt REACTOME enrichment

Supplementary Table S3.docx : Differentially methylated CpGs clusters (A-clustering regional analysis)

Acknowledgements

We wish to thank all the volunteers who participated in the project, making this study possible.

We would also like to thank Prof. S. Horvath (Department of Human Genetics, UCLA School of Medicine, Los Angeles, CA, USA) for his valuable help in Methylation Age analyses and for fruitful discussion.

Ethics statement

Our study complies with the Declaration of Helsinki principles, and conforms to ethical requirements. All volunteers signed an informed consent form at enrollment. The study protocol was approved by the Ethics Committee of the Italian Institute for Genomic Medicine (formerly Human Genetics Foundation, Turin, Italy).

Authors' contribution

GM, CM, ID, DM, SG, CV, GC conceived and designed the study

CM, DF, DM enrolled the subjects, organized the case-control study, managed personal information databases

CC, FG, RL, EP provided the MPM samples and data

DM developed the exposure assessment protocol and evaluated exposure levels for all the subjects

SG, AAllione, EC carried out the laboratory analyses

CV, GC carried out the statistical analyses

SG, AAllione, CV, EC substantially contributed to the interpretation of results

SG wrote the manuscript

GM, DM, CM, ID, MB, DF, AAspesi critically revised the manuscript and provided important intellectual content

All authors read and approved the final manuscript

References

1. Magnani C, Bianchi C, Chellini E, et al. III Italian Consensus Conference on Malignant Mesothelioma of the Pleura. Epidemiology, Public Health and Occupational Medicine related issues. *Med Lav* 2015;106:325-332.
2. Nowak AK. Chemotherapy for malignant pleural mesothelioma: a review of current management and a look to the future. *Ann Cardiothorac Surg* 2012;1:508-515.
3. Dong Y, Zhao H, Li H, et al. DNA methylation as an early diagnostic marker of cancer (Review). *Biomed Rep* 2014;2:326-330.
4. Kanherkar RR, Bhatia-Dey N, Csoka AB. Epigenetics across the human lifespan. *Front Cell Dev Biol* 2014;2:49.
5. Shivapurkar N, Gazdar AF. DNA methylation based biomarkers in non-invasive cancer screening. *Curr Mol Med* 2010;10:123-132.
6. Vandermeers F, Neelature Sriramareddy S, Costa C, et al. The role of epigenetics in malignant pleural mesothelioma. *Lung Cancer* 2013;81:311-318.
7. Dianzani I, Gibello L, Biava A, et al. Polymorphisms in DNA repair genes as risk factors for asbestos-related malignant mesothelioma in a general population study. *Mutat Res* 2006;599:124-134.
8. Betti M, Ferrante D, Padoan M, et al. XRCC1 and ERCC1 variants modify malignant mesothelioma risk: a case-control study. *Mutat Res* 2011;708:11-20.
9. Matullo G, Guarrera S, Betti M, et al. Genetic variants associated with increased risk of malignant pleural mesothelioma: a genome-wide association study. *PLoS One* 2013;8:e61253.
10. Ferrante D, Mirabelli D, Tunesi S, et al. Pleural mesothelioma and occupational and non-occupational asbestos exposure: a case-control study with quantitative risk assessment. *Occup Environ Med* 2016;73:147-153.
11. R-CoreTeam. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Available at Vienna, Austria, 2014 http://www.R-project.org/.
12. Horvath S, Levine AJ. HIV-1 Infection Accelerates Age According to the Epigenetic Clock. *J Infect Dis* 2015;212:1563-1573.
13. Allione A, Guarrera S, Russo A, et al. Inter-individual variation in nucleotide excision repair pathway is modulated by non-synonymous polymorphisms in ERCC4 and MBD4 genes. *Mutat Res* 2013;751-752:49-54.
14. Chen BH, Marioni RE, Colicino E, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging (Albany NY)* 2016;8:1844-1865.
15. Cadby G, Mukherjee S, Musk AW, et al. A genome-wide association study for malignant mesothelioma risk. *Lung Cancer* 2013;82:1-8.
16. Nishimura Y, Kumagai-Takei N, Matsuzaki H, et al. Functional Alteration of Natural Killer Cells and Cytotoxic T Lymphocytes upon Asbestos Exposure and in Malignant Mesothelioma Patients. *Biomed Res Int* 2015;2015:238431.
17. Maeda M, Nishimura Y, Kumagai N, et al. Dysregulation of the immune system caused by silica and asbestos. *J Immunotoxicol* 2010;7:268-278.
18. Miura Y, Nishimura Y, Katsuyama H, et al. Involvement of IL-10 and Bcl-2 in resistance against an asbestos-induced apoptosis of T cells. *Apoptosis* 2006;11:1825-1835.
19. Perna L, Zhang Y, Mons U, et al. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. *Clin Epigenetics* 2016;8:64.
20. Levine ME, Hosgood HD, Chen B, et al. DNA methylation age of blood predicts future onset of lung cancer in the women's health initiative. *Aging (Albany NY)* 2015;7:690-700.
21. Zheng Y, Joyce BT, Colicino E, et al. Blood Epigenetic Age may Predict Cancer Incidence and Mortality. *EBioMedicine* 2016;5:68-73.
22. Wang W, Li X, Lee M, et al. FOXKs promote Wnt/beta-catenin signaling by translocating DVL into the nucleus. *Dev Cell* 2015;32:707-718.

23. Wu M, Wang J, Tang W, et al. FOXK1 interaction with FHL2 promotes proliferation, invasion and metastasis in colorectal cancer. *Oncogenesis* 2016;5:e271.
24. Peng Y, Zhang P, Huang X, et al. Direct regulation of FOXK1 by C-jun promotes proliferation, invasion and metastasis in gastric cancer cells. *Cell Death Dis* 2016;7:e2480.
25. Nakatsumi H, Matsumoto M, Nakayama KI. Noncanonical Pathway for Regulation of CCL2 Expression by an mTORC1-FOXK1 Axis Promotes Recruitment of Tumor-Associated Macrophages. *Cell Rep* 2017;21:2471-2486.
26. Yu H, Mashtalir N, Daou S, et al. The ubiquitin carboxyl hydrolase BAP1 forms a ternary complex with YY1 and HCF-1 and is a critical regulator of gene expression. *Mol Cell Biol* 2010;30:5071-5085.
27. Testa JR, Cheung M, Pei J, et al. Germline BAP1 mutations predispose to malignant mesothelioma. *Nature genetics* 2011;43:1022-1025.
28. Betti M, Aspesi A, Biasi A, et al. CDKN2A and BAP1 germline mutations predispose to melanoma and mesothelioma. *Cancer Lett* 2016;378:120-130.
29. Betti M, Casalone E, Ferrante D, et al. Inference on germline BAP1 mutations and asbestos exposure from the analysis of familial and sporadic mesothelioma in a high-risk area. *Genes, chromosomes & cancer* 2015;54:51-62.
30. Behrends C, Sowa ME, Gygi SP, et al. Network organization of the human autophagy system. *Nature* 2010;466:68-76.
31. Darash-Yahana M, Gillespie JW, Hewitt SM, et al. The chemokine CXCL16 and its receptor, CXCR6, as markers and promoters of inflammation-associated cancers. *PLoS One* 2009;4:e6695.
32. Leone P, Shin EC, Perosa F, et al. MHC class I antigen processing and presenting machinery: organization, function, and defects in tumor cells. *J Natl Cancer Inst* 2013;105:1172-1187.
33. Keppler D. Towards novel anti-cancer strategies based on cystatin function. *Cancer Lett* 2006;235:159-176.
34. Kim H, Phung Y, Ho M. Changes in global gene expression associated with 3D structure of tumors: an ex vivo matrix-free mesothelioma spheroid model. *PLoS One* 2012;7:e39556.
35. Qi F, Okimoto G, Jube S, et al. Continuous exposure to chrysotile asbestos can cause transformation of human mesothelial cells via HMGB1 and TNF-alpha signaling. *Am J Pathol* 2013;183:1654-1666.
36. Wolff H, Vehmas T, Oksa P, et al. Asbestos, asbestosis, and cancer, the Helsinki criteria for diagnosis and attribution 2014: recommendations. *Scand J Work Environ Health* 2015;41:5-15.

Table 1: Sample characteristics and descriptive

	TRAINING SET			TEST SET		
	Cases (N=82)	**Controls (N=68)**	***p*** [a]	**Cases (N=81)**	**Controls (N=69)**	***p*** [a]
Casale M. [N (%)]	51 (62.20)	51 (75.00)	ns	49 (60.49)	43 (62.32)	ns
Torino [N (%)]	31 (37.80)	17 (25.00)		32 (39.51)	26 (37.68)	
Males [N (%)]	60 (73.17)	43 (63.24)	ns	53 (65.43)	57 (82.61)	**0.03**
Females [N (%)]	22 (26.83)	25 (36.76)		28 (34.57)	12 (17.39)	
Age [years, mean±sd]	68.46±11.41	65.16±10.07	**0.02**	66.71±10.67	64.02±10.00	ns
Exposure dose [mean±sd][b]	34.8±165.39	5.11±12.46	**2.4×10^{-6}**	62.98±194.51	2.13±3.43	**1.5×10^{-8}**
Exposure source						
Occupational [N (%)]	52 (63.41)	39 (57.35)	ns	54 (66.67)	32 (46.38)	**0.004**
Domestic [N (%)][c]	13 (15.85)	8 (11.76)		12 (14.81)	7 (10.14)	
Environmental [N (%)]	17 (20.73)	21 (30.88)		15 (18.52)	30 (43.48)	
Smoking habits						
Current smokers [N (%)]	18 (21.95)	13 (19.12)	ns	11 (13.58)	17 (24.64)	**0.02**
Former smokers [N (%)]	30 (36.59)	29 (42.65)		24 (29.63)	31 (44.93)	
Never smokers [N (%)]	34 (41.46)	26 (38.24)		41 (50.62)	21 (30.43)	
N/A	-	-		5 (6.17)	-	
Histologic Subtype						
Epithelioid [N (%)]	54 (65.85)			55 (67.90)		
Sarcomatoid [N (%)]	10 (12.20)			6 (7.41)		
Biphasic [N (%)]	12 (14.63)			16 (19.75)		
Undefined [N (%)]	1 (1.22)			2 (2.47)		
Not known [N (%)]	5 (6.10)			2 (2.47)		
Estimated WBCs percentages [d]						
CD8T % [mean±sd]	0.03±0.05	0.06±0.05	**6.3×10^{-5}**	0.03±0.03	0.07±0.04	**4.0×10^{-10}**
CD4T % [mean±sd]	0.07±0.05	0.15±0.07	**2.2×10^{-10}**	0.08±0.05	0.14±0.06	**2.3×10^{-7}**
NK % [mean±sd]	0.05±0.05	0.08±0.05	**1.3×10^{-3}**	0.06±0.04	0.07±0.05	**0.04**

B cells % [mean±sd]	0.06±0.02	0.09±0.03	**4.1×10^{-10}**	0.06±0.03	0.08±0.03	**5.0×10^{-5}**
Monocytes % [mean±sd]	0.07±0.04	0.07±0.04	ns	0.08±0.05	0.07±0.03	**0.01**
Granulocytes % [mean±sd]	0.74±0.13	0.61±0.11	**2.0×10^{-10}**	0.73±0.11	0.62±0.11	**2.6×10^{-8}**

[a] Chi-square test (two-sided, 2X2 contingency table) for Centre, Sex, Exposure source, and Smoking habits; Wilcoxon Rank Sum test (two-sided, Cases *vs* Controls) for Age, Exposure, WBC estimated percentages. Statistically significant when $p \leq 0.05$

[b] fibers/mL year

[c] Sharing house with an asbestos worker, or having asbestos-containing tools/materials in the house

[d] Percentages estimated from methylation data for all WBCs

Table 2: Reports of asbestos-induced potential immune system deregulation

Observation	References
– asbestos exposure is associated with autoimmune response – possible relationship between immunological pathways and processes leading to asbestos-related diseases – asbestos and asbestos-related inflammation may modify cellular and molecular features of immunocompetent cells eventually leading to reduction of tumor immunity	Pfau JC, *et al*, Environmental health perspectives. 2005; 113: 25-30 Noonan CW, *et al*, Environmental health perspectives. 2006; 114: 1243-7 Matsuzaki, H., *et al*, Journal of biomedicine & biotechnology, 2012. 2012: p. 492608
– significant reduction in surface expression levels of NK-cell specific activating receptors in NK cell line YT-A1 exposed to chrysotile asbestos – reduction of the intracellular serine protease granzyme A secreted by NK cells against tumor cells – reduction of degranulation events	Nishimura Y, *et al*, International journal of immunopathology and pharmacology. 2009; 22: 579-90 Nishimura, Y., *et al*, Biomed Res Int, 2015. 2015: p. 238431 Maeda M, *et al*, Journal of immunotoxicology. 2010; 7: 268-78
– MT-2 cells (human T-lymphocyte immortalized polyclonal T-lymphocyte line) exposed to short-term high dose of chrysotile asbestos showed apoptosis, ROS production, activation of the mitochondrial apoptotic pathway, similarly to what found in alveolar epithelial and pleural mesothelial cells exposed *in vitro* to asbestos – long-term low-dose exposure of the same cell line showed resistance to asbestos, long term survival, and T-lymphocyte Vβ receptors activation, suggesting an asbestos-related super-antigenic activity in human T-lymphocytes	Maeda M, *et al*, Journal of immunotoxicology. 2010; 7: 268-78 Miura Y, *et al*, Apoptosis : an international journal on programmed cell death. 2006; 11: 1825-35

– a particular fraction of CD4+ T lymphocytes acquires apoptosis resistance through activation of STAT3 and transcriptional activation of bcl-2, after direct interaction with asbestos fibers – CD4+ T lymphocytes from asbestosis patients without signs of a malignant tumor showed relative expression of bcl-2 similar to that of healthy donors – CD4+ T lymphocytes from malignant mesothelioma patients exhibited significant up-regulation of bcl-2 expression, suggesting the onset of MPM might be related to the enhanced survival of a specific population of CD4+T lymphocytes	Miura Y, et al, Apoptosis : an international journal on programmed cell death. 2006; 11: 1825-35

Table 3

Table 3: Differentially methylated CpGs in the Training set and Test set

List of the top CpGs hypomethylated with effect size ≤-0.15 and hypermethylated with effect size ≥0.10 in both Training set and Test set (cases *vs* controls)

					TRAINING-SET					TEST-SET					
					Beta-values					Beta-values					
Probe ID	**Chr**	**Map position (GRCh37/hg19)**	**Gene Symbol**	**UCSC Refgene Group**	**Cases**	**Controls**	**Eff. size±se**	**p**	**$p_{fdr\ EWAS}$**	**Cases**	**Controls**	**Eff. size±se**	**p**	**$p_{fdr\ EWAS}$**	**$p_{fdr\ TARGETED}$**
					Mean±sd	**Mean±sd**				**Mean±sd**	**Mean±sd**				
CpGs hypomethylated in MPM cases															
cg04572930	7	4754834	*FOXK1*	Body	0.35±0.17	0.54±0.16	-0.18±0.03	5.19E-08	2.46E-05	0.37±0.19	0.51±0.16	-0.17±0.03	1.19E-06	1.75E-04	2.93E-06
cg04739200	6	135517046	*MYB*	Body	0.32±0.14	0.48±0.15	-0.16±0.03	5.94E-08	2.64E-05	0.30±0.14	0.48±0.17	-0.16±0.03	5.56E-07	1.05E-04	1.62E-06
cg01521397	20	60590872	*TAF4*	Body	0.33±0.16	0.50±0.16	-0.16±0.03	5.18E-07	8.58E-05	0.35±0.15	0.51±0.14	-0.15±0.03	2.09E-07	5.85E-05	8.33E-07
CpGs hypermethylated in MPM cases															
cg08450017	3	45984838	*CXCR6; FYCO1*	TSS200 Body	0.76±0.11	0.63±0.13	0.13±0.02	1.04E-08	1.26E-05	0.73±0.13	0.63±0.11	0.10±0.02	1.39E-05	9.69E-04	2.31E-05
cg26033526	6	32819858	*TAP1*	Body	0.68±0.11	0.56±0.10	0.11±0.02	5.21E-08	2.46E-05	0.67±0.11	0.57±0.09	0.10±0.02	1.46E-06	2.00E-04	3.48E-06
cg23825480	22	31336785	*MORC2*	Body	0.78±0.10	0.66±0.10	0.10±0.02	1.28E-08	1.38E-05	0.76±0.11	0.65±0.13	0.10±0.02	2.20E-05	1.34E-03	3.43E-05
cg00446123	20	62367888	*LIME1*	TSS200	0.76±0.11	0.64±0.11	0.10±0.02	2.13E-07	5.20E-05	0.73±0.11	0.63±0.11	0.10±0.02	2.64E-06	3.03E-04	5.66E-06

Table 4

Table 4: DMRs around DM-CpGs in the Training set and Test set

CpG sites in cluster	Sites in Cluster	Map position 1st CpG in cluster (GRCh37/hg19)	Gene Symbol	UCSC Refgene Group	TRAINING-SET					TEST-SET				
					Average Beta-values		Eff. size±se	p	p_{fdr}	Average Beta-values		Eff. size±se	p	p_{fdr}
					Cases Mean±sd	Controls Mean±sd				Cases Mean±sd	Controls Mean±sd			
cg26548834; cg04261496; cg01509853; *cg26177213*; cg26136772; ***cg04572930***; cg15200418; *cg18276112*; *cg25344401*	9	Chr7:4752951	*FOXK1*	Body	0.41±0.18	0.52±0.16	-0.10±0.02	6.90E-11	2.17E-08	0.42±0.18	0.52±0.15	-0.10±0.01	7.24E-12	4.63E-09
cg26288715; cg27434890; **cg04739200**; cg18835596*	4	Chr6:135506834	*MYB*	Body	0.39±0.18	0.48±0.16	-0.09±0.02	2.20E-08	7.98E-07	0.31±0.15	0.44±0.16	-0.12±0.02	9.99E-10	7.53E-08
cg05705212; **cg08450017**; cg25226014; cg01178899	4	Chr3:45984743	*CXCR6* *FYCO1*	TSS1500 Body	0.87±0.10	0.80±0.13	0.06±0.01	6.15E-11	2.02E-08	0.86±0.11	0.80±0.13	0.05±0.01	7.15E-08	1.81E-06
cg23228341[a]; ***cg26033526***; cg01673307; cg24111025; cg02181920; cg06473288; cg17626301; cg26234900; cg10666909; cg08818207	10	Chr6:32818477	*TAP1*	Body	0.79±0.13	0.74±0.13	0.06±0.01	2.50E-08	8.75E-07	0.78±0.13	0.72±0.13	0.06±0.01	2.04E-08	6.85E-07
cg14977069; cg06653796; cg21201401; **cg00446123**; cg20513976; cg24631526; cg14396214; cg12413156	8	Chr20:62367698	*LIME1*	TSS1500	0.52±0.23	0.45±0.20	0.07±0.01	2.68E-09	1.89E-07	0.49±0.22	0.44±0.20	0.05±0.01	2.47E-05	2.10E-04

In bold the index CpG (i.e. among the 7 top DM-CpGs)
In italic underlined CpGs identified also at single-CpGs analysis (i.e. among the 868 DM-CpGs)
[a]this CpG is member of the cluster only in Training set

Table 5

Table 5: Correlation between DNA methylation and Gene Expression in the healthy pleura dataset

Probe ID	Chromosome position (GRCh37/hg19)	UCSC Refgene Group	CpG identified in Analysis	Correlation rho	Correlation p[a]	Average Beta value Mean±sd	Average Expr. value[b] Mean±sd
FOXK1[c]							
cg26548834	Chr7: 4752951	Body	Single-CpG/A-clustering	0.50	**0.02**	0.59±0.10	7.88±0.14
cg04261496	Chr7: 4753002	Body	A-clustering	0.43	**0.05**	0.50±0.10	
cg01509853	Chr7: 4754502	Body	A-clustering	0.51	**0.02**	0.87±0.04	
cg26177213	Chr7: 4754566	Body	Single-CpG/A-clustering	0.43	**0.05**	0.68±0.10	
cg26136772	Chr7: 4754681	Body	A-clustering	0.50	**0.02**	0.81±0.08	
cg04572930	Chr7: 4754834	Body	Single-CpG/A-clustering	0.50	**0.02**	0.73±0.14	
cg15200418	Chr7: 4755010	Body	A-clustering	0.47	**0.03**	0.56±0.09	
cg18276112	Chr7: 4755032	Body	Single-CpG/A-clustering	0.35	0.12	0.56±0.09	
cg25344401	Chr7: 4755415	Body	Single-CpG/A-clustering	0.58	**0.01**	0.75±0.11	
FYCO1[d]							
cg08450017	Chr3: 45984838	Body	Single-CpG	0.43	**0.05**	0.81±0.08	8.22±0.20

[a]In bold significant p-values ($p \leq 0.05$)
[b]Transcript level defined as $\log_2$ of the fluorescent signal intensity for the BeadChip probe
[c]Entrez_Gene_ID 221937, Illumina Probe ID 2082244, RefSeq ID NM_001037165.1, Protein Product NP_001032242.1
[d]Entrez_Gene_ID 79443, Illumina Probe ID 1709032, RefSeq ID NM_024513.1, Protein Product NP_078789.1

Figure 1

Click here to access/download;Figure;GUARRERA - MPM_Methylation - Figure 1.tiff

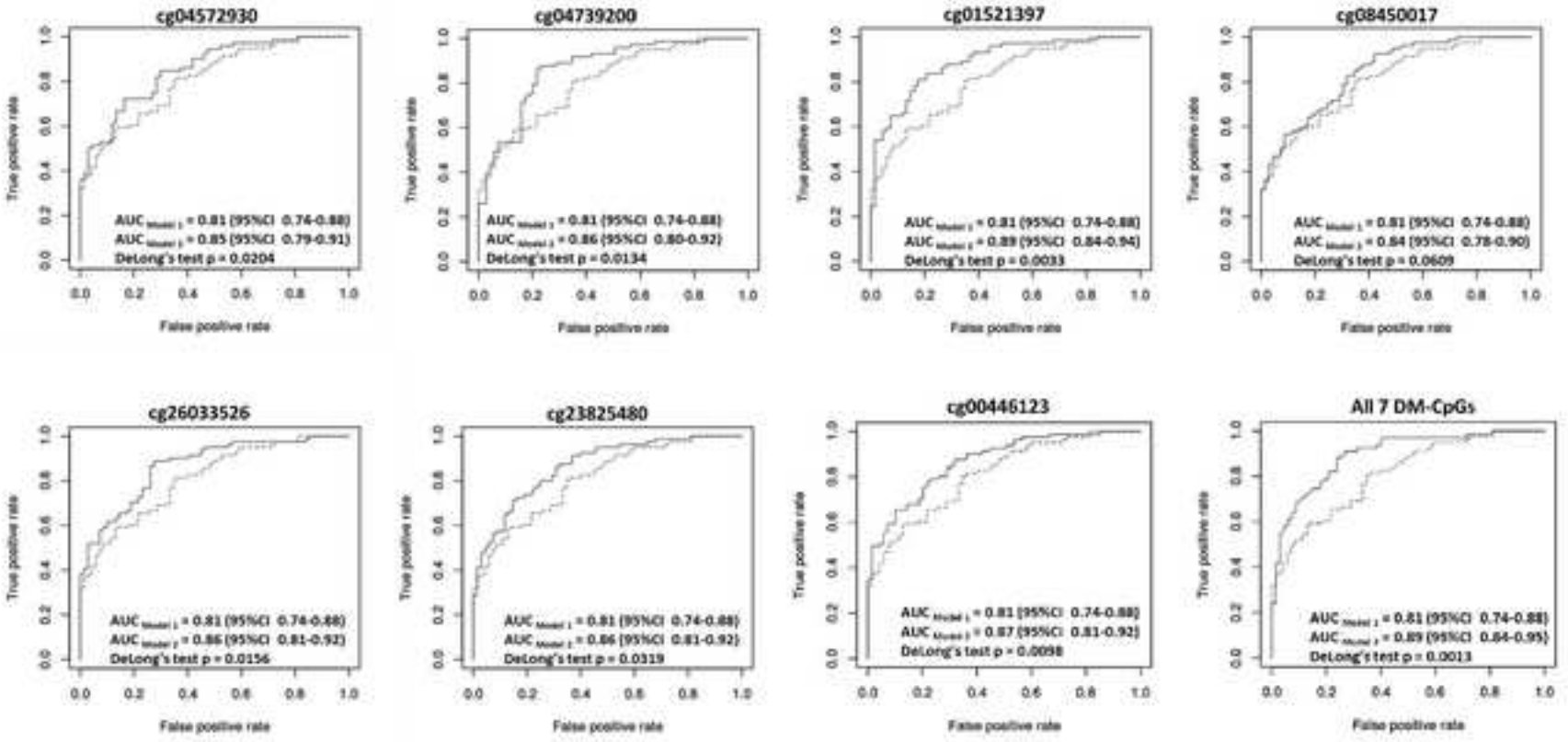
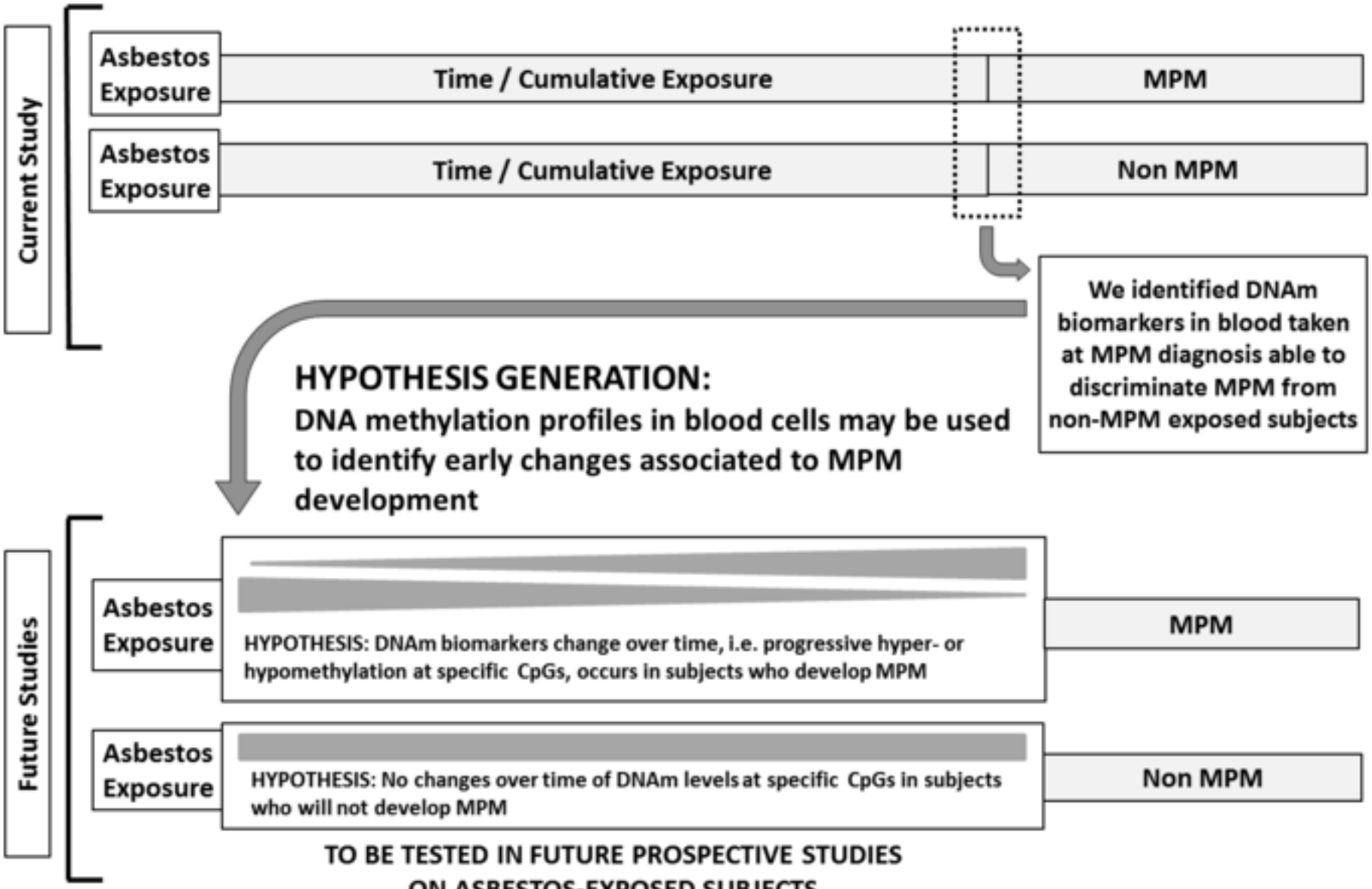


Figure 2



Supplementary Figure S1

Supplementary Figure S2